# 基于改进 FasterRCNN 的安全帽佩戴检测研究 *

徐守坤 [1†]，王雅如 [1]，顾玉宛 [1]，李　宁 [1,2]，庄丽华 [1]，石　林 [1]

(1.常州大学 信息科学与工程学院 数理学院，江苏 常州 213164；2. 闽江学院 福建省信息处理与智能控制重点实验室，福州 350108)

**摘　要：**针对已有安全帽佩戴检测算法对部分遮挡、尺寸不一和小目标存在检测难度大，准确率低的问题，提出了基于改进的 Faster RCNN 和多部件结合的安全帽佩戴检测方法。在原始 Faster RCNN 上运用多尺度训练和增加锚点数量增强网络检测不同尺寸目标的鲁棒性，并引入防止正负样本不均衡的在线困难样本挖掘策略，然后对检测出的佩戴安全帽工人和安全帽等采用多部件结合方法剔除误检目标。实验表明，相比于原始 Faster RCNN，检测准确率提高了 7%，对环境的适应性更强。

**关键词：**安全帽佩戴检测；Faster RCNN；多尺度训练；在线困难样本挖掘；多部件结合

**中图分类号：**TP391.41　　doi: 10.19734/j.issn.1001-3695.2018.07.0667

## Safety helmet wearing detection study based on improved faster RCNN

Xu Shoukun[1†], Wang Yaru[1], Gu Yuwan[1], Li Ning[1, 2], Zhuang Lihua[1], Shi Lin[1]

(1. *School of Information Science & Engineering*, *School of Mathematics & Physics*, *Changzhou University*, *Changzhou Jiangsu* 213164, *China*; 2. *Fujian Provincial Key Laboratory of Information Processing & Intelligent Control*，*Minjiang College*, *Fuzhou* 350108, *China*)

**Abstract:** Concerning the problem that the existing safety helmet wearing detection method has certain detection difficulty and low accuracy for partial occlusion, different size and small object, proposing a safety helmet wearing detection method based on improved Faster RCNN and multi-part model. Based on original Faster RCNN algorithm, this method uses multi-scale training and increased anchor number strategy to enhance the robustness of network to detect different scales, and introduce online hard example mining strategy to prevent imbalance between positive and negative samples. Then, detecting the wearing safety helmet workers, safety helmet and other parts. Last, using multi-part model method to eliminate the false object. Experiments show that the detection accuracy is improved by 7% compared to the original Faster RCNN, and the adaptability to the environment is stronger.

**Key words:** safety helmet wearing detection; faster rcnn; multi-scale training; hard negative sample mining; multi-part model

## 0　引言

化工厂、变电站和建筑工地等场所环境复杂，存在各种威胁人身安全的危险因素。由于头部是人体最关键的部位，也最容易受到致命的伤害，所以，在此类场所作业的人员均有佩戴安全帽的要求。在监控系统中自动检测工人是否佩戴安全帽并作出相应的反馈，对保护国家的劳动力，提高安全绩效等至关重要。

安全帽佩戴检测作为目标检测的一种，有着重要的应用价值。贾俊苏等人[1]使用可变形部件模型作为特征载体，将图像的几何形状特征、纹理特征和颜色特征进行组合来进行安全帽的佩戴检测。该方法设计过程非常复杂且特征的计算量较大。Kalantarikhandani 等人[2]将整个检测过程分为两部分内容，首先，结合图像的频域信息与方向梯度直方图（histogram of oriented gradient，HOG）用于工人检测；然后结合颜色和圆环霍夫变换（circle Hough transform，CHT）特征用于安全帽检测。该方法取得了一定的检测效果，但是方法整体的准确率较低且只能检测特定颜色的安全帽，以上都是基于传统的机器学习的方法进行分类检测的，这类方法多是基于人为设计选取的特征，需要实验者具有坚实的专业基础和丰富的实验经验，特征设计过程主观性较强且十分复杂，不但费时费力，而且泛化能力较差，难以适应光照等条件的变化。

随着近年来深度学习相关技术的快速发展，越来越多的研究人员将深度学习方法应用于图像分类[3]、目标识别[4]、图像分割和检测[5]等众多复杂的任务中，并且都取得了相当好的效果。基于深度学习的目标检测算法主要分为两类：一类是 RCNN 系列的基于区域的目标检测算法，如 Fast RCNN[6]、Faster RCNN[7]以及 R-FCN[8]等，这类算法的检测结果精度较高，但是速度较慢；另一类是以 YOLO[9]为代表的将检测转换为回归问题求解，如 YOLO、SSD[10]等，这类算法检测速度较快，但是精度较低且对于小目标的目标检测

**收稿日期：**2018-07-29；**修回日期：**2018-09-10　**基金项目：**国家自然科学基金应急管理项目（61640211）；2016 年度省第五期“333 工程”科研资助项目

**作者简介：**徐守坤（1972-），男（通信作者），吉林蛟河人，教授，博士，主要研究方向为人工智能、普适计算等（16106406@smail.cczu.edu.cn）；王雅如（1993-），女，安徽滁州人，硕士研究生，主要研究方向为图像处理；顾玉宛（1982-），女，江苏吴江人，讲师，博士，主要研究方向为图像处理；李宁（1974-），男，副教授，博士，主要研究方向为数据与信息处理；庄丽华（1980-），女，硕士，讲师，主要研究方向为大数据处理，图像识别；石林（1979-），男，副教授，硕士，主要研究方向为数据处理、图像识别.

效果不理想。相比基于候选区域的方法，直接预测边界框的方法能提高目标检测系统的检测速度。但 YOLO 网络直接对原图像进行网格划分，会使目标位置过于粗糙。SSD 加以改进，对不同深度的网络层回归采用不同尺度的窗口，但是由于 SSD 采用的候选框选取机制，对小目标的检测效果仍然差于 Faster-RCNN。

Faster RCNN 将目标检测实现模块（候选区域生成、特征提取、目标分类、位置精修）统一到一个深度网络框架之中，完全实现端到端的目标检测，检测准确率相对于其他深度网络都要高。因此，本文选取 Faster RCNN 网络作为安全帽佩戴检测的基础框架。但是，如果直接将原始的 Faster RCNN 模型应用于实际场景中的安全帽佩戴检测中可能会有三个弊端：a)不同场景中的佩戴安全帽工人的尺寸差异较大，可能无法检测远处的小目标;b）佩戴安全帽工人目标在图片中较背景区域所占比例小，致使 RPN 产生较多的负样本空间，网络模型难以收敛;c）由于安全帽区域占佩戴安全帽工人区域比例相对较小，网络多会将未佩戴安全帽的人员误检为目标。

基于上述原因，本文提出一种以 Faster RCNN 为基础模型，结合多尺度训练、增加锚点数量和在线困难样本挖掘（online hard example mining，OHEM）[11]机制优化原始模型，解决了目标尺寸差异大、遮挡等因素的影响，克服了训练时负样本空间过大等缺点。同时对检测出的佩戴安全帽工人及其部件（安全帽、工人）采用多部件结合方法，进一步提高了安全帽佩戴检测的准确率。

## 1　安全帽佩戴检测方法

传统的图像检测方法是基于可变形部件模型[11]（DPM）的，在 VOC2007 数据集中，它的平均精度（mean average precision，mAP）[12]可以达到 43%。研究人员将 CNN 强大的分类能力应用于图像检测。最新提出的 Faster RCNN 框架在 VOC2007 数据集上的 mAP 可达到 73%。研究表明，Faster RCNN 网络模型在各种复杂背景下对各种类别目标都具有很强的鉴别能力[7]。所以，本文将 Faster RCNN 模型应用到化工厂、变电站和施工场地等场所的安全帽佩戴检测中。

### 1.1　Faster RCNN 简介

Faster RCNN 是由两个模块组成：生成候选区域的区域建议网络（region proposal network，RPN）模块和 Fast RCNN 目标检测模块。RPN 模块产生候选区域，并利用“注意力”机制，让 Fast RCNN 有方向性的去检测目标。首先，RPN 网络预先产生可能是佩戴安全帽工人和安全帽等部件的一系列目标候选框，然后 Fast RCNN 基于提取出的候选框来对目标检测识别。

#### 1.1.1 区域建议网络（RPN）

RPN [7]的基本思想是在特征图上找到所有可能的目标候选区域，它通过在原始的网络结构上添加卷积层和全连接层来同时在每个位置上回归目标边界框和预测目标分数。RPN 网络结构如图 1 所示。RPN 采用的是滑动窗口机制，每个滑动窗口都会产生一个短的特征向量来输入到全连接层中进行位置和类别的判断，在每个滑动窗口位置同时预测多个候选区域，其中每个位置的预测候选区域的数量为 k。因此，回归层具有 4k 个输出，编码 k 个框的四个坐标，分类层输出 2k 个评分，预测每个区域的所属目标的概率和所属背景的概率。k 个候选框被参数化，称为 k 个锚点。为了使算法更好地适用于不同形状和尺寸的目标，网络将特征图上的每个位置设置多个锚点，以预测输入图像的不同尺度和不同长宽比的候选区域。

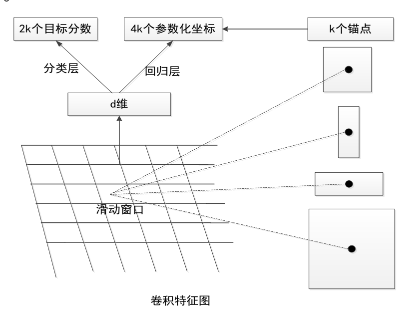



图 1　RPN 网络结构

Fig. 1　RPN network structure

#### 1.1.2 区域建议网络的损失函数

在训练 RPN 网络时，为每个候选框分配一个二值标签，用于网络训练，将以下两种情况分配正标签：

a）与某个真实目标区域框的 IoU（intersection-over-union）最大的候选框。

b）与任意真实目标区域框的 IoU 大于 0.7 的候选框。

为所有真实目标候选框的 IoU 小于 0.3 的候选框分配负标签，然后进行网络训练并微调参数。图像的损失函数定义[7]如式（1）所示。

$$L(\{p_i\},\{t_i\})=\frac{1}{N_{cls}}\sum_i L_{cls}(p_i,p_i^*)+\lambda\frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i,t_i^*) \tag{1}$$

其中：$i$ 表示小批次处理中的第 $i$ 个候选框索引，$p_i$ 是第 i 个候选框为目标的概率，若 $i$ 为候选目标，则 $p_i^*$ 为 1，否则为 0。$t_i=\{t_x,t_y,t_w,t_h\}$ 是一个向量，表示预测的参数化的候选框坐标。$t_i^*$ 是对应的是真实目标框的坐标向量。$t_i$ 和 $t_i^*$ 的定义如式（2）所示。

$$\begin{cases} t_x=(x-x_a)/w_a, t_y=(y-y_a)/h_a \\ t_w=\log(w/w_a), t_h=\log(h/h_a) \\ t_x^*=(x^*-x_a)/w_a, t_y^*=(y^*-y_a)/h_a \\ t_w^*=\log(w^*/w_a), t_h^*=\log(h/h_a) \end{cases} \tag{2}$$

其中：（$x,y$）为包围盒的中心点坐标；（$x_a,y_a$）为候选框的坐标；（$x^*,y^*$）为真实区域的包围盒的坐标，$w$ 和 $h$ 为包围盒的宽和高。算法的目的在于找到一种关系将原始框 P 映射到与真实框 G 更接近的回归框。

分类的损失函数 $L_{cls}$ 定义如式（3）所示。

$$L_{cls}(p_i,p_i^*)=-\log[p_i^* p_i+(1-p_i^*)(1-p_i)] \tag{3}$$

回归的损失函数 $L_{reg}$ 定义如式（4）所示。

$$L_{reg}(t_i,t_i^*)=R(t_i-t_i^*) \tag{4}$$

其中 R 是 smooth $_{L1}$ 函数，smooth $_{L1}$ 函数如式（5）所示。

$$smooth_{L1}(x)=\begin{cases} 0.5x^2, & |x|<1 \\ |x|\text{-}0.5, & |x|\ge 1 \end{cases} \tag{5}$$

#### 1.1.3 RPN 与 Fast RCNN

Fast RCNN 是基于 RPN 产生的候选区域边界框来检测并识别候选框中的佩戴安全帽工人。训练过程中，首先将带有标注的数据集输入到网络中，通过卷积网络进行特征提取，然后将 RPN 产生对应的候选区域映射到该特征图中，获得相应的特征信息。再通过 ROI 池化层产生固定大小的特征图输入到全连接层中，产生目标所属类别概率和区域坐标。RPN 和 Fast RCNN 通过交替优化共享特征提取网络部分，实现高效的目标检测。

### 1.2　改进的 Faster RCNN

为了使 Faster RCNN 模型能够更好的应用于安全帽佩戴检测，本文提出在原始 Faster RCNN 基础上结合多尺度训练、增加锚点数量和在线困难样本挖掘机制策略增强模型的性

能。

1.2.1 多尺度训练

在实际施工现场，佩戴安全帽工人的目标尺寸差异较大。原始的 Faster RCNN 通常会对所有的训练图片采用固定尺寸，这样对不同尺寸的目标检测泛化性能较差。本文采用多尺度训练，在将图片送入网络之前，在保证图像原有比例的前提下，将图片随机调整大小，使其较短边取 480，600 和 750 之一。然后随机选择三种尺度之一送入网络中进行训练。实验证明多尺度训练使得网络能够学习目标的各种尺寸的特征，使得网络对目标尺寸大小具有一定的鲁棒性。

1.2.2 锚点数量

RPN 网络中的锚点数量是网络中一个非常重要的超参数，它直接影响到后续的候选区域的生成。原始的 Faster RCNN 使用 9 个锚点，每个滑动窗口产生 9 个不同尺度和不同长宽比的候选区域，最后对整张图片产生的候选区域使用非极大值抑制算法剔除多余的候选区域。但是网络默认设置的锚点参数对区域较小的目标无法召回。在对施工现场安全帽佩戴检测时，需要检测出安全帽、人员等多个目标，小目标总是存在的。因此，在默认参数的基础上，加入一组 64×64 锚点（比默认设置的更小）使得网络可以检测到更多的小目标。在训练过程中，RPN 部分使用 12 个锚点，尺度大小分别为 64×64，128×128，256×256，512×512，三个纵横比分别为 1：1，1：2 和 2：1。实验证明本文增加的 64×64 的尺度可以检测更多的小目标。

1.2.3 在线困难样本挖掘策略

Faster RCNN 在进行模型训练时，RPN 网络会随机产生大量的候选区域，由于目标在图片中所占比例较小，导致产生的负样本空间过大，正样本和负样本的存在巨大的失衡，网络模型偏向负样本。原始的 Faster RCNN 将 mini-batch 设置为 2，产生 128 个候选区域，输入到后续网络中进行分类检测。实际上，RPN 网络产生的候选区域超过 128 个，这些候选区域是从 RPN 产生的所有候选区域中随机选择的。正例和负例之间的数量比通常为 1：3[14]，主要用来解决训练时正负样本数量的极大失衡问题，这本质上并没有根除样本不均衡问题，且随机挑选的样本中包含大量无效样本，导致模型退化。

为了解决训练过程中正负样本失衡问题，本文将在线困难样本挖掘机制引入到 Faster RCNN 网络框架中来进行安全帽佩戴检测，在不需要人为设定样本的正负比例和不降低网络实时性能的情况下，同时进行困难样本挑选，从而提网络性能。

困难样本就是网络在进行分类时错误目标被分类为正确的且置信度阈值较高的样本。普通的困难样本选择方法是对使用一定比例的正负样本进行分类器训练，然后使用测试集进行检测，挑选困难样本加入到初始训练集中再重新进行训练，但是每次重新训练时都要固定模型，难以在神经网络中实现。Shrivastava 等人[15]提出了针对 Fast RCNN 网络模型的在线困难样本挖掘机制。本文将 OHEM 机制引入 Faster RCNN 网络框架中解决网络训练过程中负样本空间过大问题。引入 OHEM 机制的 Faster RCNN 框架如图 2 所示。

将 Faster R-CNN 的 ROI 池化层后面部分称为 ROI 网络，引入 OHEM 机制后将原始的 1 个 ROI 网络扩展成 2 个共享网络参数的 ROI 网络。其中一个是只读的，只读 ROI 网络中的所有操作都是前向的，主要功能包括根据式（1）计算所有候选区域的 loss 值并对其进行排序，选择部分 loss 值较大的候选区域，即选择当前网络表现较差的的候选区域作为困难样本。另一个 ROI 网络（图中阴影部分）包含前向和后向操作，它的输入是前一个 ROI 网络所挑选出的困难样本，输出是预测的分类结果和边界框的坐标。即首先使用一个额外的 ROI 选取困难样本，然后使用这些困难样本来进行标准 ROI 网络训练。该算法不需要设置正负样本之间的数量比例来解决数量不平衡问题，更有针对性，进一步提高了物体检测的准确性。通过实验发现，在线困难样本挖掘策略可以增强算法的判别能力，提高网络检测精度。

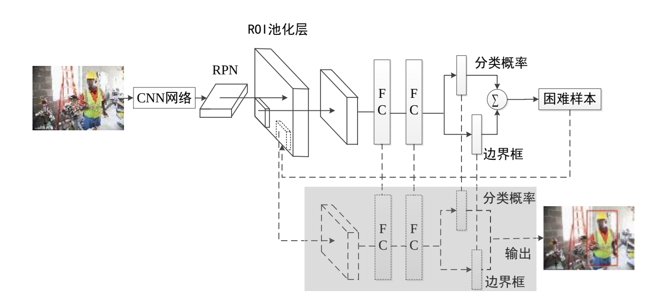



图 2　改进后的 Faster RCNN 网络框架

Fig. 2　Improved Faster RCNN network framework

## 1.3　多部件结合

佩戴安全帽的工人是由安全帽和工人两个部分组成，对图像中的各部分进行标定，然后使用改进的 Faster RCNN 框架进行训练。由于安全帽区域占佩戴安全帽工人区域比例相对较小，网络多会将未佩戴安全帽的工人误检为目标。而在检测过程中，如果有一个区域被检测为佩戴安全帽的工人，而且此区域内又存在相应位置的安全帽等部件，那么此区域检测为佩戴安全帽工人的概率就大大增加。针对此特点，提出了多部件结合的方法，用于佩戴安全帽工人检测。整个方法流程如图 3 所示。Faster RCNN 部分主要是进行数据集格式转换和模型优化等过程；多部件结合技术用来判断佩戴安全帽工人目标区域内的相应位置是否存在所属部件，从而提高检测准确性。

方法具体过程如下：

a）通过从 VOC2012 数据集挑选[16]、网络搜集和自行采集等方式获取的各种背景场景和不同质量的变电站以及施工场地等场所的监控图片，并对图片中的安全帽、人员、佩戴安全帽工人标注出来转换为 VOC2007 数据集格式。

b)将处理好的数据集输入到改进的 Faster RCNN 模型训练。

c）降低模型检测置信度阈值，然后基于上述优化的模型来检测佩戴安全帽工人以及安全帽等部件，剔除孤立部件。

d)对于剩下的佩戴安全帽工人目标，计算有无所属部件。

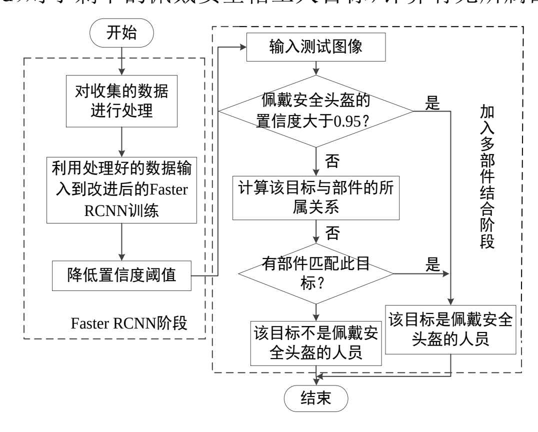



图 3　安全帽佩戴检测方法流程

Fig. 3　Method flow of safety helmets wearing detection

在使用改进的 Faster RCNN 框架进行佩戴安全帽工人等初步检测后，将检测目标分为两类：佩戴安全帽工人和安全帽等部件。若佩戴安全帽工人的检测置信度小于 0.95[17]，则通过计算部件与佩戴安全帽工人的相对位置关系以及重叠率来判断它们的从属关系，部件与目标重叠率最大且相对位置关系正确（例如，安全帽在佩戴安全帽工人区域的上部 1/3 内）的目标即判定为佩戴安全帽工人。重叠率公式[18]如（6）所示：

$$IoU = \frac{PartAcreage \cap OverallAcreage}{PartAcreage \cup OverallAcreage} \tag{6}$$

其中:*PartAcreage* 是检测后框出安全帽等部件区域，*OverallAcreage* 是框出的佩戴安全帽工人区域，*IoU* 是部件与目标的重叠率。接着，将孤立的检测结果剔除，剩下的就是希望检测到的佩戴安全帽工人。

## 2 实验分析

实验环境配置：GPU：GeForce GTX 1080Ti，CUDA8.0，Ubuntu16.04，内存 12GB。实验使用基于 CAFFE（convolution architecture for feature extraction）学习框架进行相关代码和参数训练，网络框架 Faster RCNN 选择 VGG 16[19]网络，VGG16 网络相对于 ZFNet[20]和 LeNet[21]拥有更深的网络层，更能提取目标的显著特征，取得更好的检测结果。

### 2.1 数据集

由于关于安全帽佩戴检测研究未曾发现公共数据集，所以本实验使用的数据是通过从 VOC2012 数据集中收集、自行采集和网上收集等方式获得的，共 7000 张，包含各种背景场景和不同质量的施工场地和变电站等场所的监控图片，根据实验要求，将其转换成 VOC2007 数据集格式，如图 4 所示，对各个部件进行人工标注。另外搜集了 200 张实际施工场景中的监控图片来进行方法测试，对提出方法的有效性进行验证。



图 4 标注图片数据

Fig. 4 Label image data

### 2.2 实验结果和分析

为了评估提出方法对佩戴安全帽工人检测的有效性，实验使用精确度（precision）和召回率（recall）[22]来衡量方法的有效性，计算公式如式（7）（8）所示。

$$precision = \frac{TP}{TP + FP} \tag{7}$$

$$recall = \frac{TP}{TP + FN} \tag{8}$$

其中：*TP*（true positive）表示被模型预测为正值的正样本；FP（false positive）表示被模型预测为负值的正样本；FN（false negative）表示被模型预测为负值的正样本。

实验中网络结构部分参数如表 1 所示。

表 1 VGG16 网络结构的部分参数表

Table 1 Partial parameter list of VGG16 network structure

| 类型/层数 | 卷积核数量 | 卷积核大小/步长 |
|---|---|---|
| Conv1_x/2 | 64 | 3×3/1 |
| Maxpool | | 2×2/2 |
| Conv2_x/2 | 128 | 3×3/1 |
| Maxpool | | 2×2/2 |
| Conv3_x/3 | 256 | 3×3/1 |
| Maxpool | | 2×2/2 |
| Conv4_x/3 | 512 | 3×3/1 |
| Maxpool | | 2×2/2 |
| Conv5_x/3 | 512 | 3×3/1 |

本文将网络中输出大小相同的卷积层归为一组，如表 1 中第 1 列所示。整个网络分为五组卷积层，每组包含 x 卷积层，如 Conv1_x/2，表示第一组包含 2 个卷积层。表中第 2、3 列分别表示卷积核数量和卷积核大小/步长。

1）对比相同场景下 Faster RCNN 改进前后的检测效果

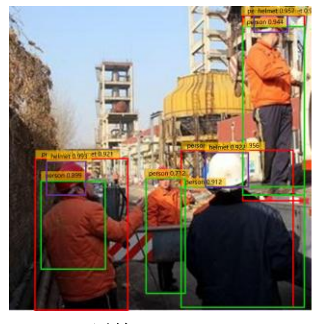
使用 7000 张 VOC2007 格式的数据分别对原始的 Faster RCNN 网络和结合多尺度训练、增加锚点数量和在线困难样本挖掘机制的 Faster RNN 训练，使用 200 张实际场景中的监控图片（包含 377 个佩戴安全帽工人目标）进行训练好的两个模型进行测试。两种算法效果如表 2 所示。

表 2 两种模型效果对比

Table 2 Comparisons between two models

| 检测方法 | 检测 | 错检 | 漏检 | 召回率/% | 精确度/% |
|---|---|---|---|---|---|
| Faster RCNN | 310 | 44 | 67 | 82.22 | 87.57 |
| 改进后的 Faster RCNN | 341 | 32 | 36 | 90.45 | 91.42 |

表 2 显示，改进后 Faster RCNN 的测试精确度提高了 3.85%，召回率提高了 8.23%。图 5 是两种算法在实际图片上的检测效果，绿色框框出的是人员，紫色框框出的是安全帽，红色框框出的是佩戴安全帽工人。从图 5 中可以看出，改进后的 Faster RCNN 检测效果明显优于原始 Faster RCNN，图 b 相对于图 a 能够检测出较多的佩戴安全帽工人，且检测出的置信度也较高。实验证明，改进后的 Faster RCNN 框架能够有效地优化模型。



(a)原始 faster RCNN

(a)Original faster RCNN



(b)改进后的 faster RCNN

(b)Improved faster RCNN

图 5 两种算法的检测效果对比

Fig. 5 Comparisons with two different methods

2) 采用不同策略训练网络时的效果

为了进一步验证本文提出方法的有效性，分别使用不同策略对网络进行训练和测试，效果如表 3 所示。

表 3　使用不同策略的网络测试效果

Table 3　Network test results using different strategies

| 策略 | 锚点框数目 | 多尺度训练 | 在线困难样本挖掘 | 召回率/% | 精确度/% |
|---|---|---|---|---|---|
| 1 | 9 | 否 | 否 | 82.22 | 87.57 |
| 2 | 12 | 否 | 否 | 83.96 | 88.36 |
| 3 | 12 | 是 | 否 | 86.31 | 89.51 |
| 4 | 12 | 否 | 是 | 87.59 | 89.88 |
| 5 | 12 | 是 | 是 | 90.45 | 91.42 |

从表 3 中可以看出不同策略对网络检测性能的影响。策略 1 和策略 2 相比使得检测精确度提升了 0.79%，这是因为本文增加一组（64*64）尺度的锚点，将锚点从原始的 9 个增加到 12 个，使得网络能够检测到更多的小目标。策略 2 与策略 3 相比网络模型的测试精度提升了 1.15%，这是因为网络模型在训练阶段采用多尺度训练的策略，使得网络模型对不同尺寸的目标具有一定的鲁棒性。对比策略 2 和策略 4 检测精确度提升了 1.52%，这是因为本文将在线困难样本挖掘机制嵌入到网络模型中，解决了网络在训练过程中的负样本空间过大问题，增强了网络的分辨力。实验证明，三种策略均对提升了网络模型的检测性能。

3）置信度阈值确定

为了在初始检测阶段获得较多的佩戴安全帽工人和安全帽等目标和部件，需在 Faster RCNN 的初检阶段设置一个较低的置信度阈值。实验基于改进后的 Faster RCNN 框架对模型检测置信度阈值讨论。表 4 列出了不同置信度阈值下的目标检测效果。

表 4　不同置信度阈值下目标检测效果

Table 4　Target detection effect under different confidence thresholds

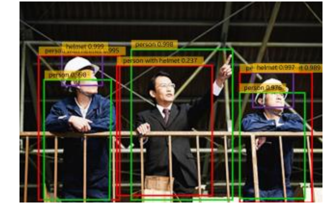
| 置信度阈值 | 错检率/% | 漏检率/% |
|---|---|---|
| 0.1 | 19.1 | 3.4 |
| 0.2 | 16.5 | 3.4 |
| 0.3 | 12.6 | 7.5 |
| 0.4 | 8.3 | 13.3 |

如表 4 所示，当置信度阈值为 0.2 时，佩戴安全帽的人员的漏检率最低，而错检率相对阈值为 0.1 时也有所降低，为了最大程度的提高检测的准确性，本实验选取置信度阈值为 0.2。如图 6 所示，当置信度阈值选为 0.2 时，图 a 中有一个置信度 0.854 的佩戴安全帽工人被检测出来，并且有较低置信度的所属部件也被检测出，使得初始检测阶段可以检测出较多目标和部件。相对地，也会有错误的目标被检测出，如图 b，中间人员被误检为佩戴安全帽工人。此时需要对检测出的目标进行筛选。



(a)检出低置信度目标

(a)Detect low confidence targets



(b)存在误检目标

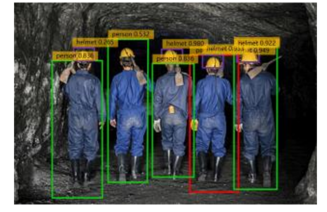
(b)Existing false detection targets

图 6　部件初检结果

Fig. 6　Initial detection result

4）对比使用多部件结合方法

在降低检测置信度阈值之后，多部件结合的方法进一步检测图片中佩戴安全帽的人员，使用收集到的 200 张现实施工现场的图片进行方法测试，检测效果如图 7 所示。



(a)存在漏检

(a) Existing missed targets



(b)降低置信度阈值检测结果

(b) Detection result by reducing the confidence threshold




(c)存在误检目标

(c) Existing false detection targets




(d)多部件结合方法剔除误检目标

(d) Multi-parts method to eliminate false detection targets

图 7　多部件结合效果

Fig. 7　Results of multi-part model

采用多部件结合的方法，通过计算出部件和目标重叠率和相对位置关系正确，判断此目标为佩戴安全帽工人，从而降低了漏检率。在降低置信度阈值之后，图 c 误检出一个目标为佩戴安全帽的人员，通过多部件结合的方法，判断出所属部件的相对位置关系不正确（安全帽的位置），故将此目标剔除，如图 d。在进行多部件结合方法之后，去除掉部件检测框，得到最终检测目标。除了原始的 Faster RCNN 模型，本文还与 YOLO 模型以及 HOG 特征方法进行实验对比，对比效果如表 5 所示，可以看出本文提出方法检测效果明显，测试精确度达到 94.23%。

表 5　对比多部件检测效果

Table 5　Detection results compared with multi-part model

| 检测方法 | 召回率/% | 精确度/% |
|---|---|---|
| HOG+SVM | 71.95 | 79.83 |
| YOLO | 80.86 | 85.43 |
| 改进 Faster RCNN | 90.45 | 91.42 |
| 本文方法 | 91.51 | 94.56 |

5）不同场景和不同质量的检测效果

测试集包含不同场景和质量的图片，如图 8 所示，网络对目标尺寸差距较大如图 8（a）所示、检测目标在光线较差的环境下如图 8（b）所示、目标存在部分遮挡以及检测目标较为模糊如图 8（c）（d）的场景中均有较好的检测效果，验证了方法的有效性。

## 3　结束语

本文首先使用多尺度训练、增加锚点数量和在线困难样本挖掘策略优化原始的 Faster RCNN 网络模型；然后将从

VOC2012 数据集和自行采集等方式收集到的数据输入到改进的 Faster RCNN 中进行模型训练，提取出佩戴安全帽工人和安全帽等部件；最后采用多部件结合的方法来进一步提高安全帽的佩戴检测精度。实验结果证明，本文使用的方法对安全帽的佩戴检测精度有显著提高，对多尺寸目标和小目标均有较好的检测效果，并对场景的变化和部分遮挡也有很好的鲁棒性。然而，因为人员的姿态不一，而此方法只能粗略的选取安全帽等部件的相对位置。在以后的研究中，将着重这方面进行展开，以求达到更好的泛化性。



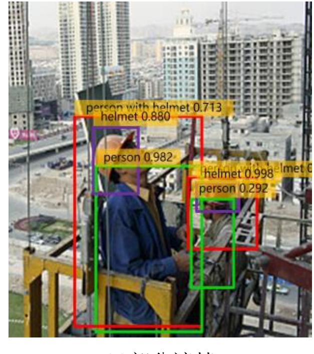
(a)目标尺寸差距大

(a) Large target size gap



(b)光线较差

(b) Poor light




(c)部分遮挡

(c) Partial occlusion




(d)目标模糊

(d) Target blur

图 8 不同场景下的安全帽佩戴检测效果

Fig. 8 Safety helmets wearing detection results in different scenes

## 参考文献：

[1] 贾峻苏，鲍庆洁，唐慧明. 基于可变形部件模型的安全头盔佩戴检测 [J]. 计算机应用研究，2016, 33(3): 953-956. (Jia Junsu, Bao Qingjie, Tang Huiming. Method for detecting safety helmet based on deformable part model [J]. Application Research of Computers, 2016, 33(3): 953-956. )

[2] Rubaiyat A H M, Toma T T, Kalantari-Khandani M, *et al.* Automatic detection of helmet uses for construction safety [C]//Proc of International Conference on Web Intelligence Workshops. IEEE Press, 2017: 135-142.

[3] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]//Proc of International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012: 1097-1105.

[4] Donahue J, Jia Yangqing, Vinyals O, *et al.* DeCAF: a deep convolutional activation feature for generic visual recognition [C]// Proc of International Conference on International Conference on Machine Learning. 2014: I-647.

[5] Druzhkov P N, Kustikova V D. A survey of deep learning methods and software tools for image classification and object detection [J]. Pattern Recognition & Image Analysis, 2016, 26 (1): 9-15.

[6] Girshick R. Fast R-CNN [C]// Proc of IEEE International Conference on Computer Vision. IEEE Press, 2015: 1440-1448.

[7] Ren S, He Kaiming, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137-1149.

[8] Nong Song, Ni Zihan. Gesture recognition based on R-FCN in complex scenes [J]. Huazhong Keji Daxue Xuebao, 2017, 45(10): 54-58.

[9] Redmon J, Divvala S, Girshick R, *et al.* You only look once: unified, real-time object detection [C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2016: 779-788.

[10] Liu Wei, Anguelov D, Erhan D, *et al.* SSD: single shot multibox detector [C]//Proc of European Conference on Computer Vision. Springer International Publishing, 2016: 21-37.

[11] Chu Jun, Guo Zhixian, Leng Lu. Object detection based on multi-layer convolution feature fusion and online hard example mining [J]. IEEE Access, 2018, PP (99): 1-1.

[12] Felzenszwalb P F, Girshick R B, Mcallester D, *et al.* Object detection with discriminatively trained part-based models [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2010, 32(9): 1627.

[13] Ceri S, Bozzon A, Brambilla M, *et al.* An introduction to information retrieval [J]. Pharmacogenomics Journal, 2013, 2(2): 96-102.

[14] Ren Yun, Zhu Changren, Xiao Shunping. Object detection based on fast//faster rcnn employing fully convolutional architectures [J]. Mathematical Problems in Engineering, 2018, 2018 (1): 1-7.

[15] Shrivastava A, Gupta A, Girshick R. Training region-based object detectors with online hard example mining [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2016: 761-769.

[16] Everingham M, Gool L V, Williams C K I, *et al.* The pascal visual object classes (VOC) challenge [J]. International Journal of Computer Vision, 2010, 88(2): 303-338.

[17] Yates F. The Influence of statistical methods for research workers on the development of the science of statistics[J]. Journal **of the American Statistical Association,** 2012, 46(253): 19-34.

[18] Xia Yizhang, Zhang Bailing, Frans Coenen. Face Occlusion Detection Using Deep Convolutional Neural Networks [J]. International Journal of Pattern Recognition & Artificial Intelligence, 2016, 30(9): 401-408.

[19] Li Jianjun, Peng Kangjian, Chang Chinchen. An Efficient Object Detection Algorithm Based on Compressed Networks [J]. Symmetry, 2018, 10(7): 235.

[20] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks [C]// Proc of European conference on Computer Vision,. Cham: Springer International Publishing AG, 2014: 818-833.

[21] Haykin S, Kosko B. Gradient based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[22] Byeon Y H, Kwak K C. A Performance Comparison of Pedestrian Detection Using Faster RCNN and ACF [C]// Proc of IIAI International Congress on Advanced Applied Informatics.Washington DC:IEEE Computer Society, 2017: 858-863.